description, but by the claims and their equivalents, and all variations within the scope of the claims and their equivalents are to be construed as being included in the disclosure.

What is claimed is:

1. An apparatus for normalizing input data of an acoustic model, the apparatus comprising:

a window extractor configured to extract windows of frame data to be input to the acoustic model from frame data of a speech to be recognized; and

a normalizer configured to normalize the frame data to be input to the acoustic model in units of the extracted windows.

2. The apparatus of claim **1**, wherein the window extractor is further configured to consecutively extract the windows in units of a predetermined number of frames of the frame data of the speech to be recognized while the frame data of the speech to be recognized is being input.

3. The apparatus of claim **1**, wherein the normalizer is further configured to normalize frames belonging to a current window together with padding frames added to both sides of the current window.

4. The apparatus of claim **1**, wherein the normalizer is further configured to normalize frames belonging to a current window in consideration of frames belonging to preceding windows of the current window.

5. The apparatus of claim **4**, wherein the normalizer is further configured to normalize the frames belonging to the current window in consideration of the frames belonging to the preceding windows and frames of training data in response to a total number of the frames belonging to the current window and the frames belonging to the preceding windows being insufficient for speech recognition.

6. The apparatus of claim **5**, wherein the normalizer is further configured to acquire a number of frames corresponding to a difference between the total number of the frames and a reference value from the training data in response to the total number of the frames being less than the reference value.

7. The apparatus of claim **1**, wherein the normalizer is further configured to normalize the frame data belonging to the extracted windows so that the frame data belonging to the extracted windows has an average of 0 and a standard deviation of 1.

8. A method of normalizing input data of an acoustic model, the method comprising:

extracting windows of frame data to be input to the acoustic model from frame data of a speech to be recognized; and

normalizing the frame data to be input to the acoustic model in units of the extracted windows.

9. The method of claim **8**, wherein the extracting of the windows comprises consecutively extracting the windows in units of a predetermined number of frames of the frame data of the speech to be recognized while the frame data of the speech to be recognized is being input.

10. The method of claim **8**, wherein the normalizing of the frame data comprises normalizing frames belonging to a current window together with padding frames added to both sides of the current window.

11. The method of claim **8**, wherein the normalizing of the frame data comprises normalizing frames belonging to a current window in consideration of frames belonging to preceding windows of the current window.

12. The method of claim **11**, the normalizing of the frame data comprises normalizing the frames belonging to the current window in consideration of the frames belonging to the preceding windows and frames of training data in response to a total number of the frames belonging to the current window and the frames belonging to the preceding windows being insufficient for speech recognition.

13. The method of claim **12**, wherein the normalizing of the frame data comprises:

comparing the total number of the frames belonging to the current window and the preceding windows with a reference value in response to the current window being extracted; and

acquiring a number of frames corresponding to a difference between the total number of the frames and the reference value from the training data in response to the total number of the frames being less than the reference value.

14. The method of claim **8**, wherein the normalizing of the frame data comprises normalizing the frame data belonging to the extracted windows so that the frame data belonging to the extracted windows has an average of 0 and a standard deviation of 1.

15. A non-transitory computer-readable medium storing instructions that, when executed by a processor, cause the processor to perform the method of claim **8**.

16. A speech recognition apparatus comprising:

a preprocessor configured to:

extract windows of frame data to be input to an acoustic model from frame data of a speech to be recognized, and

normalize the frame data to be input to the acoustic model in units of the extracted windows;

an acoustic score calculator configured to calculate acoustic scores in units of the normalized windows using the acoustic model based on a deep neural network (DNN); and

an interpreter configured to:

interpret the acoustic scores calculated in units of the normalized windows, and

output a recognition result of the speech to be recognized based on the interpreted scores.

17. The speech recognition apparatus of claim **16**, wherein the preprocessor is further configured to normalize frames belonging to a current window in consideration of frames belonging to preceding windows of the current window.

18. The speech recognition apparatus of claim **17**, wherein the preprocessor is further configured to normalize the frames belonging to the current window in consideration of the frames belonging to the preceding windows and frames of training data in response to a total number of the frames belonging to the current window and the frames belonging to the preceding windows being insufficient for speech recognition.

19. The speech recognition apparatus of claim **16**, wherein the interpreter is further configured to output a recognition result of the current window as a final recognition result of a whole speech to be recognized in response to a predetermined condition being satisfied or an input of a user while input of the speech to be recognized is under way.

20. The speech recognition apparatus of claim **16**, wherein the DNN is a bidirectional recurrent deep neural network (BRDNN).